Swedish University of Agricultural Sciences
The Faculty of Natural Resources and Agricultural Sciences
Uppsala BioCenter
Department of Microbiology

Role of promoter strength in gene regulation in *Kcnq1* imprinted domain

Shokoufeh Karimi

Master´s thesis • 30 hec • Second cycle, A2E
Biotechnology • Examensarbete / Sveriges lantbruksuniversitet, Institutionen för mikrobiologi, 2012:1 • ISSN 1101-8151
Uppsala 2012

Role of promoter strength in gene regulation in *Kcnq1* imprinted domain

Shokoufeh Karimi

Supervisor:	Tomas Linder, Swedish University of Agricultural Sciences, Department of Microbiology
Assistant Supervisor:	Chandrasekhar Kanduri, Uppsala University, Department of immunology, genetics and pathology
Examiner:	Volkmar Passoth, Swedish University of Agricultural Sciences, Department of Microbiology

Credits: 30 hec
Level: Second cycle, A2E
Course title: Independent project in Biology
Course code: EX0565
Programme/education: Biotechnology

Place of publication: Uppsala
Year of publication: 2012
Title of series: no: Examensarbete / Sveriges lantbruksuniversitet, Institutionen för mikrobiologi, 2012:1.

ISSN: 1101-8151
Online publication: http://stud.epsilon.slu.se

Key Words: *Kcnq1* domain, Imprinting, Long non RNA, Epigenetic regulation, Promoter strength

Swedish University of Agricultural Sciences
The Faculty of Natural Resources and Agricultural Sciences
Uppsala BioCenter
Department of Microbiology

Abstract

Genomic imprinting is a genetic phenomenon by which some of the genes go through several processes that restrict them from biallelic expression. To date, many investigations have been focused on mechanism of genomic imprinting in gene clusters. Though, there have been several attempts to uncover the mechanisms behind this complex phenomenon, but it is not fully yet understood.

Kcnq1 is a genomic imprinting domain in mammalian genomes that is located on chromosome 7 in mouse and on chromosome 11p15.5 in human; consists of 8-10 paternally imprinted genes and a maternally imprinted non-coding gene known as *Kcnq1ot1* that encodes for a long non-coding RNA. Additionally, a number of non-imprinted genes are also localized between these genes in this locus. These imprinted genes are classified into two distinct groups including placenta specific genes, which are expressed only from embryonic tissue and ubiquitously imprinted genes, which expression is repressed in both placenta and embryo's tissues.

A growing body of evidence indicates several roles for long non-coding RNA *Kcnq1ot1* in imprinting of these genes in a tissue and cell type specific manner. It has been shown that this long non-coding RNA regulates the imprinting via DNA and chromatin modification at the *Kcnq1* domain.

This thesis aimed to investigate the effect of promoter strength on imprinting. The Promoter strength differences among these genes could be one of the possible factors involving differential patterns of imprinting within *Kcnq1* locus. The promoter activities of five genes including two ubiquitously (*Kcnq1* and *Slc22a18*), two placenta specific (*Ascl2* and *Tssc4*) and a non-imprinted genes (*Nap1l4*) were assayed. We showed that the non-imprinted gene and placenta specific imprinted genes have higher promoter activities than ubiquitously imprinted genes. Our data indicate that degree of promoter strength is involved in imprinting of these genes in the locus and those genes with stronger promoter can escape silencing. However, further investigations are needed to be done to define the role for promoter strength in tissue-specific imprinting.

Keywords: Kcnq1 domain, Imprinting, Long non RNA, Epigenetic regulation, Promoter strength

To My Family

CONTENTS

ABBREVATIONS

%- Percent
°C- Degree Celsius
5Mc- 5-Methylcytosine
bp- base pair
BWS- Beckwith-Wiedemann Syndrome
CO2- Carbon dioxide
CpG- Cytosine-Guanine dinucleotide
CTCF- CCCTC-Transcription Factor
d.p.c- day post coitum
DMR- Differentially Methylated Region
DNA- Deoxyribonucleic Acid
DNMT- DNA Methyl Transferase
FBS- Foetal Bovine Serum
HBS-Hank's Balaced Salt
HMT- Histone Methyltransferase
H3K27Me3- Histone H3 lysine 27 trimethylation
H3K20Me3- Histone H3 Lysine 20 trimethylation
H3K9Me3- Histone H3 Lysine 9 trimethylation
ICR- Imprinting Control Region
Igf2- Insulin-like Growth Factor 2
LCR- Locus Control Region
LncRNA- Long non-coding RNA
M- Molar
Mb- Mega base
MBD3- Methyl CpG-Binding Domain Protein three
MetS- Metabolic Syndrome
NcRNA- Non-coding RNA
PBS- Phosphate Buffered Saline
PGC- Primordial Grem Cells
PHP-Ib- Pseudohypoparathyoridism type Ib
Pmol- Pico mol
PRC-Polycom Repressor Complex
PRY- Polymerase Chain Reaction
PWS- Prader-Willi Syndrome
RNA- Ribonucleic Acid
RNAi- RNA interference
RLB- Reporter Lysis Buffer
TND-Transient Neonatal Diabetes
UTR- Untranslated Region

Epigenetic has always been all the weird and wonderful things that can not be explained by genetics. Denise Barlow (Vienna, Austria)

1 INTRODUCTION

In organisms, genetic information is stored as DNA. Though in multicellular organisms, DNA sequence is essentially the same in all the cells of one individual, cells are different in their function due to the way genes are regulated. Regulation of gene expression is an essential process and enhances the flexibility that living organisms need to cope with their own environment.

The mechanisms involved in gene regulation are complex and have been in the center of attention for decades. Several mechanisms related to gene expression have been described and categorize in different stages including chromatin core, transcription, post-transcriptional modification, RNA transport, and translation and mRNA degradation [1].

In mammals, DNA and chromatin modifications are responsible for gene regulation in DNA level and this regulation is under influence of different biological processes and signals such as transcription factor and epigenetic machinery including DNA methylation, histone modification and non-coding RNA (ncRNA) based mechanism [2]. Though, transcription factors are key regulators of gene expression, a growing body of evidences suggests the role of multilayer epigenetic mechanisms in gene expression [2]. In addition, it has been shown that the expression patterns of genes are different in male and female. Usually, most of the genes are expressed from both alleles but there are genes that are expressed just from one allele (i.e. monoalleleic expression) in a sex–specific manner, which has been described as gene imprinting [3].

The work presented in this thesis describes our attempt to understand the role of promoter strength in expression of imprinting genes.

1.1 Epigenetic modification

For the first time, the term “Epigenetics” was employed by Conrad Waddington in 1939, when he proposed the idea of epigenetic landscape to describe the process of cell differentiation and how different undifferentiated cell types have potential to develop along many paths to become a particular tissue during the multicellular organismal develop-ment [4, 5].

Currently, the term of ‘Epigenetics’ applies to study heritable mechanism of gene function without any DNA sequence’s changes [5].

Epigenetics possess very complicated mechanisms to organize different processes and signals to start and maintain the epigenetic marks [2].

However, we do not know much about epigenetic initiatory such as non-coding RNA and different proteins that bind to the DNA, epigenetic maintainers such as DNA methylation, histone modification have been well-studied [2].

1.1.1 DNA Methylation

In majority of mammals, DNA methylation is the most important epigenetic modification that is involved in regulating the imprinting. In 1948, even before discovery of DNA structure, 5-methylcytosine (5Mc) was discovered by Rollin Hotchkiss [6]. In 1975, methyltransferase has been characterized and three different families of the DNA (cytosine-5) methyltransferase (DNMT) have been identified. These families include a maintenance methyltransferase DNMT1, which causes the methylation maintenance by methylation of unmethylated DNA strand during DNA replication and de novo methyltransferases DNMT3 (DNMT3a & DNMT3b) that set up de novo methylation patterns of DNA in the early embryonic development through newly methylation of DNA [7]. Moreover, DNMT3L and DNMT2 have been discovered, though the catalytic motifs needed for methyltransferase activity have not been identified yet [7, 8].

DNA methylation generally takes place in CpG dinucleotide in CpG islands, which is initiated and maintained by different families of methyltransferases [9]. CpG islands are stretches of CpG nucleotide that normally possess three criteria including the GC content of more than 50%, more than 0.6 observed/expected ratio of CpG dinucleotide and finally they should be more than 400 base pairs in length. They are usually linked to a regulatory element. For example, 50% to 60% of the promoter regions of human genes contain CpG islands [9].

The CpG island methylation plays an essential role in gene regulation and subsequently mammalian development. The majority of CpG Islands are unmethylated in promoter regions of mammalian genome except in imprinted genes, female inactive X chromosome and de novo methylation during cell differentiation [9].

Several studies revealed that deletion of methyltransferases like DNMT1 and/or DNMT3a/DNMT3b and aberration of CpG island methylation in imprinting gene are associated with embryonic fatality in mice and many human diseases [7, 9].

1.1.2 Histone modification

In eukaryotes, very large genomic DNA is compressed to form chromatin. The central unit of chromatin is known as nucleosome, composed of 147 base pairs of nucleotides wrapped around an octamer core histone including dimers of H2A, H2B, H3 and H4 histone proteins. Usually the N-terminal domains of these proteins are protruded and they are subjected to different post transcriptional modification processes such as phosphorylation, ubiquitylation, methylation, acetylation and sumoylation [10].

In mammals, histone methylation occurs in histone H3 and H4, which is mediated by five histone methyltransferases (HMTs) including G9a, Suv39h1, Suv39h2, Eset and Eu-HMTase. All of these HMTs transfer the methyl group to lysine (k) residues of histone H3 [11].

Histone modifications mediate formation of specific structures of chromatin i.e., euchromatin and heterochromatin. These histone changes in euchromatin regions are methylations at lysine 4, 36, 79 of

H3 whereas in heterochromatin region methylations take place at lysine 9, 27 of H3 and 20 of H4 [11, 12].

1.1.3 Non-coding RNA

It was commonly believed that complex organisms had larger genomes than the simple ones. However, the term of genome paradox was introduced to explain that the relative biological complexity of the higher organisms do not correlate with the DNA content and the number of the protein coding genes [13]. For example, Human has been estimated about ~30000 gene. In comparison, *Caenorhabditis elegans* (*C- elegans*) genome codes for approximately ~19000 genes (http:/ /www.ensemble.org). Though, the *C- elegans* comprises only ~1000 cells but the number of the protein-coding genes is close to the one from human with millions of cells. Thus, it is obvious that the existence of protein-coding RNA alone is not enough to explain complexity of higher organisms [13].

In the multicellular organisms, about ~98 % of the transcripts are intronic and intergenic non-coding RNA. These biomolecules that were thought to be junk in earlier studies are essential for various vital and regulatory processes. In higher organism such as mammals, regulation of the genes are mediated by non-coding RNA rather than having more number of genes as believed earlier being the reason for their complex regulation. Thus, non-coding RNA is one of the key answers for the genome paradox [14].

Non-coding RNAs have been classified based on either the function or number of their nucleotides [14]. Based on their function, they are classified into two distinct groups of housekeeping ncRNAs and regulatory ncRNAs. The housekeeping ncRNAs are expressed always in all the stages of development such as ribosomal RNAs (rRNAs), small nuclear (snRNAs), transfer RNAs (tRNAs), snoRNAs whereas, regulatory RNAs are expressed irregularly at particular stages of development [14] such as short regulatory ncRNA (e.g., microRNAs, small interfering RNAs and Piwi ncRNA) and long regulatory ncRNA [15]. Besides, they also have been divided to three classes based on their number of nucleotides including microRNAs with 21 –25 nucleotides, small RNAs with 100 – 200 nucleotides and large RNAs up to over 10,000 nucleotides. They possess specific characteristics including a high number of stop codons, absence of an open reading frame and regulation of gene expression using special mechanisms such as imprinting, gene silencing, RNA interference and demethylation of DNA [16].

1.2 Genomic imprinting

To date, few model systems have been described to investigate role of epigenetics in gene expression including genomic imprinting, X-inactivation and metastable epialleles. Among them genomic imprinting is an interesting model to study some of these processes and signals including DNA methylation, histone modification, non-coding RNA and nucleosome location [17].

1.2.1 Imprinted genes

In the early 1980, for the first time, a specific gene regulation-gene imprinting- was described in mouse embryo by which in the early stages of development some of the genes were expressed from

one of the two parental chromosomes [18]. This achievement was due to creation of two mouse models (gynogenetic—'GG') and (androgenetic—'AG'). These two models show lethal phenotypes due to shortage or overexpression of imprinted genes [18]**.** Consequently, Insulin-like growth factor-2 receptor (Igf2r) was the first gene that was found to be imprinted [19]. To date, in mouse genome 144 genes have been identified that undergo imprinting and some of them are conserved in human genome too. The summaries of mouse imprinted genes are accessible in following website; Medical Research Council (MRC), (http://www.mousebook.org/catalog.php?catalog =imprinting).

These imprinted genes are including both protein coding and non-coding genes. Most of these imprinted genes are organized in a cluster that contains imprinted and non-imprinted genes, and some genes encoding non-transcribed RNAs [20].

The imprinted genes have common characteristics in the DNA level. They contain few introns, a greater number of CpG islands, a higher number of direct repeats within /near CpG islands, low compactness of SIN-Alu and SIN- MIR elements compared to non- imprinted genes, mono parental methylation of CpG islands and asymmetric replication timing of two alleles of imprintinted genes [21].

1.2.2 Aberration of imprinting associated with human disease

Several studies revealed the fact that imprinting genes are important in fetal and placental growth and developments. Additionally, many of imprinting genes are also involved in other processes such as energy homeostatic, neurological pathways, endocrine/paracrine pathways and cancer formation [17].

Human congenital disorders such as Silver–Russell, Beckwith–Wiedemann (BW) and Prader–Willi and Angelman syndromes, different forms of neoplasia and tumorigenesis, metabolic syndrome (MetS) [17], endocrine/metabolic disorders such as 6q24 transient neonatal diabetes (TND) [22] and Pseudohypoparathyoridism type Ib (PHP-Ib) are a few examples of imprinting disruption that associated with human disorders [23].

1.2.3 The life cycle of imprint

In early embryo, epigenetic machinery and reprogramming influence the various developmental stages. In mouse embryo, primordial germ cells (PGC) can be distinguished from other cell types at 7.25 day post coitum (d.p.c) and migrate into the genital cord between 10.5 and 11.5 d.p.c. It has been shown that epigenetic changes display between days 10.5 and 12.5 d.p.c [24]. The epigenetic modifications are started with genome wide DNA de-methylation and then continue with re-establishment of new sex-specific imprint marks i.e., histone modification and DNA methylation which finally should be maintained in somatic cells throughout the course of development [25].
In conclusion, the life cycle of the imprint can be divided to three different steps including erasure, establishment and maintenance.

1.2.4 Erasure

The first step of imprint life cycle is erasure of the earlier pattern of DNA methylation on the maternal and paternal alleles in the gametes. Several Studies revealed that the demethylation is completed by 12.5 d.p.c, whereas other researchers showed some of the primordial germ cells losing their

methylation in 10.5 or 11.5 d.p.c [24]. Thus, in mouse normally, this erasing process occurs after entering of primordial germ cells (PGCs) to the genital cord about days 10.5 to 13.5 of gestation [24].

1.2.5 Establishment and maintenance of imprint

When the DNA methylation has been erased from the former generation (i.e., the first step of the imprint life cycle has accomplished) the imprint life cycle is continued by second step (i.e., regaining monoallelic expression of imprinted genes) by re-establishment of DNA methylation in germline DMRs and somatic DMRs at both gametes and embryos of male and female [24]. In female and male germline DMRs, establishment of the genomic imprints take place at different phases of development. In mouse male germline, establishment of methylation marks in the paternal imprint begins about 14.5 d.p.c and is completed in perinatal. However, it is remained throughout haploid phase. In contrast, maternal imprinting in female is obtained asynchronously at different loci after birth during oocyte growth. Though, it is maintained in somatic cells throughout the course of development [26].

The imprinting life cycle is very complex and there are some other examples from DNA methylation that are also involved in establishment and maintenance of the imprinting.

1.2.6 Factors involved in establishment of imprint

Various factors have been observed to be important in establishment of female germline methylation, but not in the male germline. For instance, interaction between DNMT3L as a non-functional de novo methyltransferase and DNMT3a and/ DMNT3b is essential for establishment of methylation in female oocyte, whereas in male, such interaction is non-functional due to low CpG content of promoter region in paternal DMR and CpG spacing compared to its counterpart (i.e., maternal DMR) [26] .

In addition to histone modification of DMR in somatic cells such as H3K4 methylation, which seems to be necessary for recruiting the DNMT3A/DNMT3L complex [27]. Transcription through DMR (i.e., transcription in *Gnas* imprinting locus) is involved in establishment of germline DNA methylation. *Gnas* locus contains imprinting *Gnas, Nesp, Gnasxl* and non-coding *exon1A* and *Nespas* genes. This locus also comprises two DMRs in maternal allele including DMR1 and DMR2, which cover the promoter regions of *Nespas, Gnasxl* and *1A* genes, respectively. These two DMRs are methylated in maternal allele [28]. Deletion of the promoter region of *Nesp* and truncation of its transcript have been associated to interruption of methylation at the imprinting *GNASXL* and non-coding exon *1A* genes DMRs of human's maternal germline [26].

There are also indications that the non-histone proteins such as KRAB zinc finger ZFP57 and CTCFL play an essential role in establishment of methylation at the Snrpn and H19 DMR in maternal and paternal germ cells respectively [26].

The role of non-coding RNA in establishment of imprinting has been well studied and will be explained in detail in this thesis later.

1.2.7 Factors involved in maintenance of imprints

After establishment of methylation imprints, the imprints have to be maintained and transferred to all somatic cells. One of the important factors that mediates the maintenance of imprints is DNMT1 maintenance methyltransferase. Intriguingly, methylated DMRs of both paternal and maternal alleles

can escape from genome wide demethylation as a result of DNMT1 expression from both oocyte and pre implanted embryo [29].

In addition to DNMT1, there are proteins such as ZFP57 a recruiter of KAP-1/TIF1bc repressor complex, methyl CpG-binding domain protein three (MBD3) and CTCF proteins are other examples of molecular participation in DNA methylation maintenance [30]. Besides, *RBBP1* and *RBBP1-like1* are also involved in maintenance of H4K20me3 and H3K9me3 in nucleosome and DNA methylation in snrpn ICR [30].

1.3 Mechanism of genomic imprinting in cluster

Most of the imprinting genes are found to be in clusters and their expression are usually regulated by one or few differentially methylated regions (DMRs) known as imprinting control region (ICR) [20]. Mainly, two different models of regulation including insulator and long non-coding RNA models have been described to explain the role of ICR in imprinting locus [20].

1.3.1 Differentially methylated regions (DMRs)

Differentially methylated regions (DMRs) are cis acting elements that are essential for allele-specific expression of imprinted gene. The methylation of these DMRs is a significant mark for almost all the imprinted genes. They are methylated differentially in a sex -specific manner and limited to promoter regions; however few of them are in intronic region such as *Kcnq1* and *Igf2r* DMR2 [26, 30].

To date in mouse genome, 21 germline DMRs have been identified [26].

1.3.2 Insulator model of regulation

In the imprinted loci, some of the imprinted genes share same regulatory elements such as enhancers. In this model, insulator controls expressions of such imprinted genes through blocking the enhancer activity [31].
In the mouse *H19/Igf2* locus, imprinted genes (i.e., *H19* and *Igf2*) are regulated according to such a model. *H19* is maternally expressed and codes for a 2.2-kb ncRNA, which is a precursor for a micro RNA known as *miR-675*, whereas *Igf2* is paternally expressed and encodes a fetal growth factor. The *H19* and *Igf2* genes share common enhancers. The expressions of these two imprinted genes are under control of a DMD /ICR at 5' end of *H19* and CTCF (CCCTC-binding factor) proteins [30].

On the maternal alleles, CTCF binds to the unmethylated alleles of ICR1 and as a result, an insulator is formed. The insulator blocks the common enhancers to activate *Igf2* gene at 3'end of *H19*. Consequently, the enhancers have opportunity to interact with *H19* promoter and activate it. In contrast, on the paternal alleles, insulator is not formed due to DMD/ICR methylation, subsequently CTCF cannot bind and *Igf2* gene is expressed (Fig.1) [30, 32].

In addition to *H19/Igf2*, there are some other loci in the genome that may be regulated according to the insulator model for the regulation of gene expression, such as the *Kcnq1* and *Rasgrf1* loci [30].

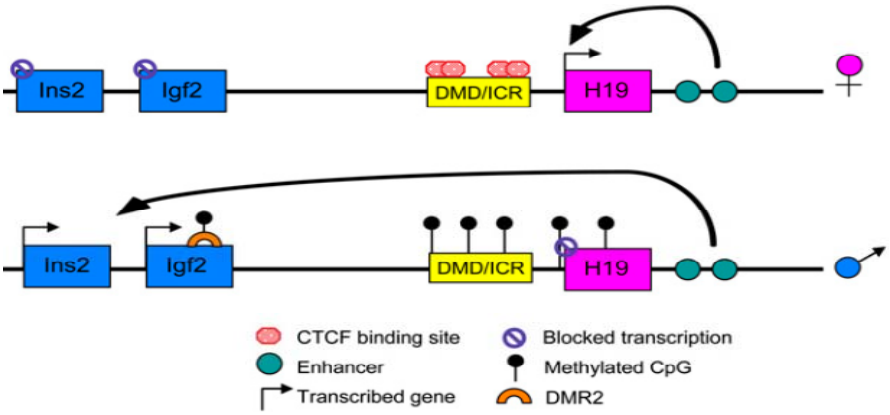


Figure 1. Mechanism of imprinting at the mouse *H19/Igf2* locus. Paternally imprinted *H19* gene encodes for an ncRNA and the maternally imprinted *Igf2* encodes a fetal growth factor that share common enhancers. On the maternal chromosome, a zinc finger protein CTCF can bind to unmethylated DMD/ICR and forms an insulator. Consequently, common enhancers are prevented to activate the *Igf2*. While enhancers are free near the *H19* promoter, as a result *H19* is expressed. On the paternal allele insulator does not form due to methylation of DMD/ICR and subsequently CTCF cannot bind and *Igf2* gene is expressed. Adopted from Bartolomei M.S. (2009).

1.3.3 Long non coding RNA model of regulation

As I mentioned above, insulator model of imprinting regulation implies for few loci, but most imprinting loci are documented for long non-coding RNA model of regulation, for example *Igf2r/Airn, Kcnq1* and *Gnas* locus [30].

The *Igf2r/Airn* imprinting locus has been used as a simplest model to explain this type of regulation. *Igf2r/Airn* locus located on mouse chromosome 17, contains three protein coding/imprinted genes (*Slc22a2, Slc22a3, Igf2r*) and one non-coding gene (*Airn*) [30]. All three protein coding genes are paternally imprinted, whereas *Airn* is maternally imprinted. In the mouse *Igf2r* locus, there are two DMRs including DMR1 and DMR2 that are located in the *Igf2r* promoter and the second intron of *Igf2r,* respectively. The promoter of Airn ncRNA is located in DMR2. The Airn antisense non-coding RNA represses the expression of the *Igf2r* as well as *Slc22a2* and *Slc22a3* from the paternal allele (Fig.2) [32, 33].The silencing role of this long non-coding RNA is due to function of RNA itself, the process of its own transcription and recruiting repressive histone marks [30].

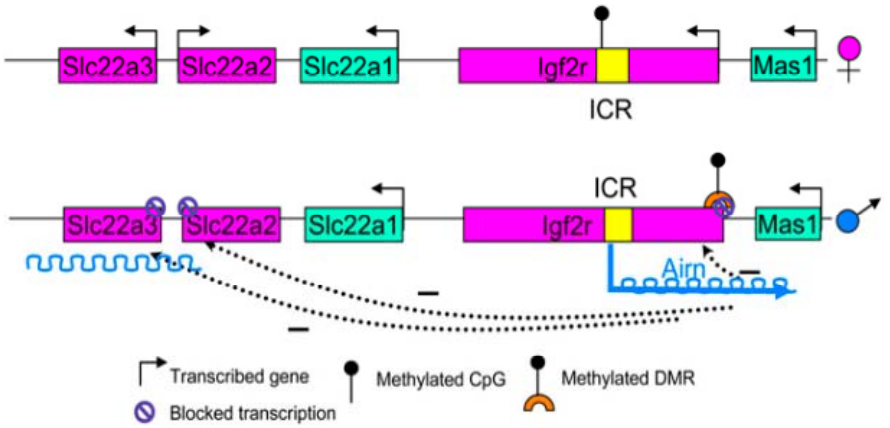


Figure 2. Imprinting mechanism across the Igf2r/Air locus. The *Igf2r, Slc22a2* and *Slc22a3* genes are paternally imprinted as a result of unmethylated ICR that is located at promoter region of Airn non-coding RNA. *Airn* and *Igf2r* are imprinted, maternally and paternally, respectively due to methylation of their promoters in an ICR. *Mas1 and Slc22a1* are non-imprinted genes. Adopted from Bartolomei. M.S (2009).

1.4 Mechanism of long non-coding RNA in gene regulation

Long non-coding RNAs are groups of non-translated transcripts that differ from housekeeping and other regulatory non-coding RNAs [15]. To date, there is no suitable definition for ncRNAs yet. However, they are similar to mRNA, but do not code for a protein and are longer than 200bp in length. They generally have a short ORF compared to protein coding genes [34].
Long non-coding RNA is involved in gene regulation through many biological processes such as DNA methylation, chromatin remodeling, nuclear architecture and subnuclear compartments formations [35].

1.4.1 LncRNA and chromatin remodeling

As I mentioned above, long non-coding RNAs are involved in gene expression by recruiting different factors. For example, during the process of X- inactivation, *RepA* long noncoding RNA along with other non-coding RNA i.e., *Xist, Tsix, Xite* are thought to be involved in establishment and maintenance of X–chromosome inactivation (XCI). *RepA* is believed to recruit the Polycomb Repressive Complex 2 (*PRC2*) directly and activates the *Xist* transcription. Consequently, maintenance of X–chromosome inactivation is due to interaction among *Xist-PRC2-RepA*, which is required for H3K27 trimethylation. This kind of mechanism has been seen in other imprinted loci such as *Kcnq1* or *Igf2/Airn* [34, 35].

Another non-coding RNA is 108 kb Air non-coding RNA that mediates the chromatin remodeling by recruiting the G9a histone methyltransferase at chromatin of *Slc22a3*'s promoter to initiate the H3K9 methylation and consequently silencing of *Slc22a3* gene in mouse placenta. [35] In the same way, *Kcnq1ot1* long non-coding RNA recruits the G9a histone methyltransferase and PRC2 in a linage specific manner [35]. *HOTAIR* ncRNA also recruits PRC2, which is necessary for H3 lysine-27 trimethylation and chromatin silencing at *HOXC* locus [34, 36].

1.4.2 LncRNAs and nuclear architecture and subnuclear compartments

The eukaryotic nucleus is well-organized, there are sub nuclear compartments including Cajal bodies, nucleoli, paraspeckles, and nuclear speckles. They are involved in several biological processes, for example paraspeckle is a ribonucleoprotein body which is associated with long non-coding RNA and involved in nuclear retention of RNA to perform particular gene regulation [37].

For instance, mouse *Ctn* RNA is an 8 kb nuclear poly (A) + RNA and was discovered as first Paraspeckle RNA involved in gene regulation [38]. The 3'UTR of *Ctn* RNA with specific inverted repetitive elements makes a RNA hairpin loop which undergoes adenosine to inosine hyper editing. Under stress conditions, the hairpin loop of the *Ctn* RNA is cleaved by a cleavage factor (CFIm) to form and regulate a shorter product *mCAT2* mRNA [37]. Another localized long non-coding RNA is *NEAT1* (*Men ε/βin* mouse) which is also involved in paraspeckle formation. It has been shown that deletion or in vitro overexpression of this gene interrupts or increases the paraspeckle formation in nucleus, respectively [35].

Some of the lnRNA have their distinct space in the nucleus such as *Xist* and *Kcnq1ot1,* which are involved in chromatin modification [35].

1.4.3 *Kcnq1* locus and *Kcnq1ot1* RNA

Kcnq1 imprinted locus is located at mouse chromosome 7 or on human chromosome 11p15.5, respectively. It contains 8-10 paternally imprinted genes including *Kcnq1, Cdkn1c, Slc22a18, Tssc4, Phlda2, Osbpl5, Ascl2, Cd81* and one maternally imprinted gene known as kcnq1ot1 (Fig.3) [39].
These genes have been classified into two main categories, the ubiquitously imprinted genes *Kcnq1, Cdkn1c, Phlda2* and *Slc22a18*, which are repressed in both embryonic and placental tissues, whereas the expression of second group, the placental-specific imprinted genes is limited to embryo tissue. These include *Cd81*, *Osbpl5, Ascl2* and *Tssc4* [40]. In addition, there are non-imprinted genes that are located among imprinted genes in the locus [39].

The allele-specific expression of imprinted genes in the kcnq1 locus is controlled by a differentially methylated region known as KvDMR/ ICR2, which is located in intron 10 of kcnq1 gene [41]. The *Kcnq1* ICR contains the promoter of kcnq1ot1 antisense RNA. When the ICR is methylated on the maternal allele, expression of *Kcnq1ot1* is repressed, whereas on the paternal allele ICR is not methylated, subsequently *Kcnq1ot1* is expressed and triggers the bidirectional silencing of the genes within the locus [41].

Kcnq1ot1 non-coding RNA is a product of RNA polymerase II with 91.5 kb length. *Kcnq1ot1* owns an 890 bp silencing domain at its 5' end. A study based on an episomal system demonstrated that deletion of this silencing domain reactivates the flanking reporter genes. The *Kcnq1* ICR contains several conserved repeated motifs including A, A1, A2 and MD1. Among them A1 and A2 are located in the 890 kb silencing domain of *Kcnq1ot1* non-coding RNA and point mutation of A2 motif decline the silencing function of *Kcnq1ot1* RNA on transcriptional level [41].

In addition, several conserved elements *CCAAT* have been identified in the *Kcnq1ot1* promoter. The CCAAT elements in the *Kcnq1ot1* promoter interact with NFY transcription factors to mediate transcription of this promoter. Mutation of these elements, obliterate the bidirectional activity of this promoter [42].

It has been demonstrated that *Kcnq1ot1* long non-coding RNA is involved in imprinting gene regulation through both chromatin and DNA modification. *Kcnq1ot1* interacts with G9a histone methyltransferase and PRC2 in a lineage-specific manner and recruits DNA methylation machinery to regulate ubiquitously imprinted genes [40]. Pandy et.al revealed that *Kcnq1ot1* does not interact likewise along the *Kcnq1* domain. *Kcnq1ot1* specifically is attracted to particular regions within the locus and has more affinity to the chromatin in placenta, but not in liver tissue [43].

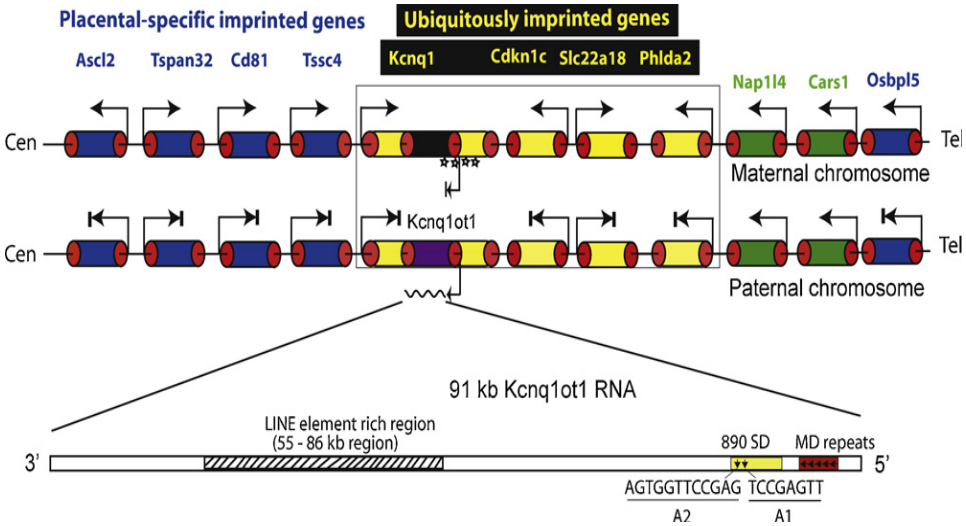


Figure 3. *Kcnq1* locus. The physical map showing 91.5 kb *Kcnq1ot1* RNA, placental-specific and ubiquitously imprinted genes. Adopted from Kanduri, C. 2011.

2 AIM AND HYPOTHESIS OF THE PROJECT

The aim of this project is to investigate the effect of promoter strength on imprinting. We expected to discriminate promoter strength differences that could explain and may uncover one of the multilayer molecular mechanisms that are involved in regulation of imprinting in *Kcnq1* locus.

In the other word, to find any link between the promoter activity and imprinting that could explain the mechanism how some genes in the *Kcnq1* imprinting domain maintain the imprinting where as others are able to escape.

3 RESULTS

3.1 Cloning of Promoter and Plasmid constructs

The research here focuses on *Kcnq1* imprinted locus, which contains eight protein coding genes /imprinted genes and a non-coding gene/imprinted gene and few non-imprinted genes. The *Kcnq1ot1* gene encodes for a 91.5 kb length long non-coding RNA, which is an antisense transcript that acts as a bi-directional silencer of neighboring genes in order to suppress their expression [41].

In this study, we have cloned the promoter region of five different genes from the *Kcnq1* imprinting locus including two placenta specific imprinted genes (*Ascl2* and *Tscc4*), two ubiquitously imprinted genes (*Kcnq1* and *Slc22a 18*) and a non-imprinted gene (*Nap1l4*) that allowed us further evaluation of the promoter activity of both imprinted and non imprinted genes in the imprinted locus (Fig 3.1).

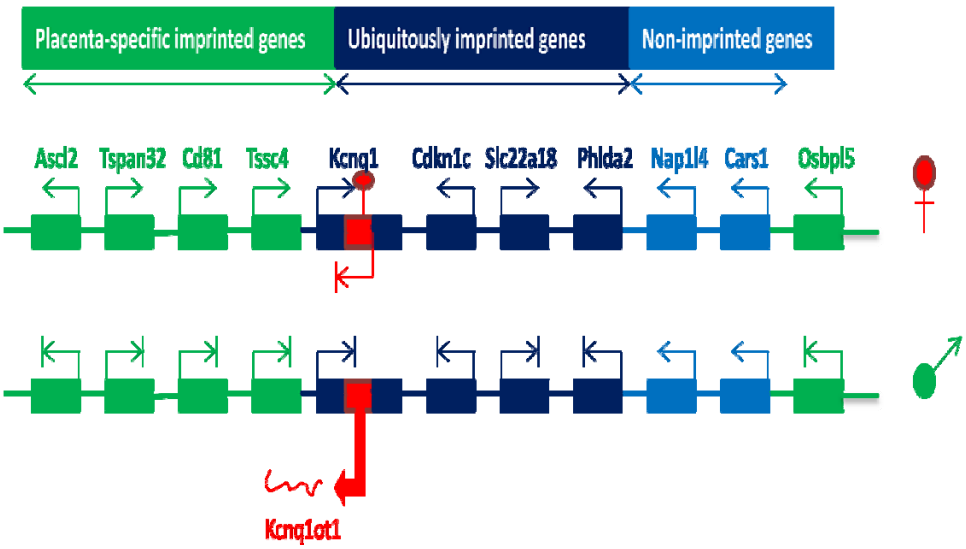


Figure 3.1. *Kcnq1* locus. The physical map showing non-imprinted, placental-specific and ubiquitously imprinted genes in *Kcnq1* imprinted domain.

To be able to perform the study of these promoters, specific PCRs were performed to amplify the DNA sequence of interest with primer binding sites using DNA extracted from mice liver and primers containing *Kpn* I (in forward primer) and *Bgl* II (in reverse primer) sites at their 5' ends. We were able to obtain PCR products for all the designed primers for the promoters. In addition, we have designed two distinct primers for one of the imprinted gene's promoter *Kcnq1* to obtain two different product sizes. The size of shorter product was 1.2 kb and longer product's size was 1.9 kb. We studied affect of the different sizes of the promoter region on gene's promoter activity and expression of reporter gene from the vectors (Fig 3.2 A, B and C).

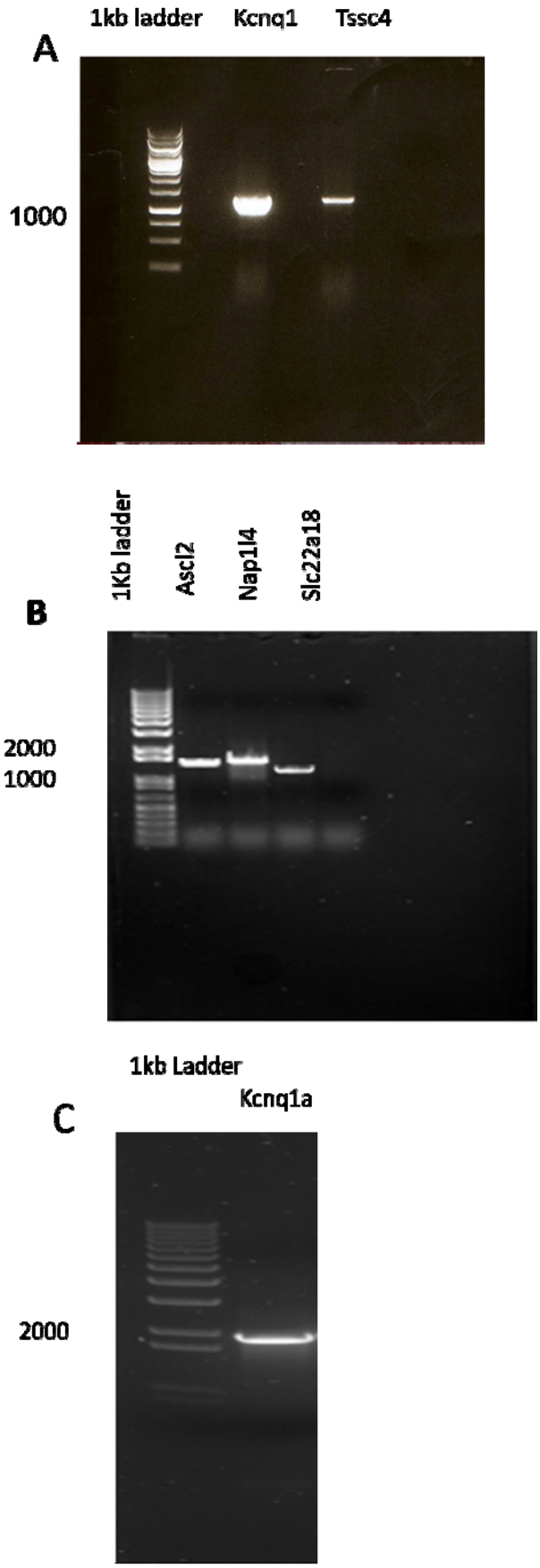
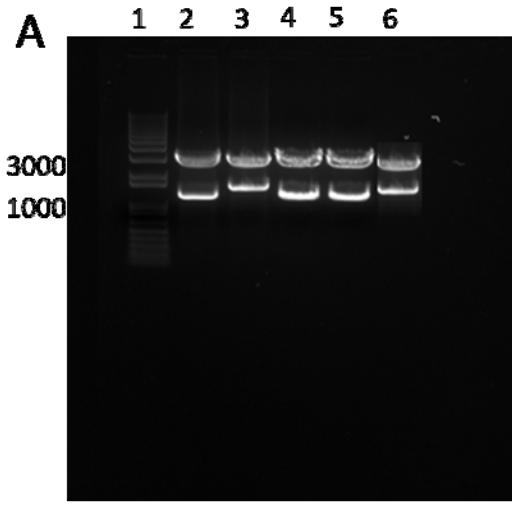


Figure 3. 2. Representative results of PCR for promoter region of five different genes including *Kcnq1* and *Tssc4* (A), *Ascl2*, *Nap1l4* and *Slc22a18* (B), and *Kcnq1a* genes (C), loaded next to the 1kb ladder. The two different sets of primer are used for *Kcnq1* gene in order to obtain two different product sizes: 1.2 and 1.9 kb.

The PCR products were extracted from the gel and cloned into the pGEM-T Easy Vector. After ligation reaction of each promoter in pGEM- T easy vector, to make sure ligation has occurred, *XL-Blue* strain of *Escherichia coli* cells were transformed with the ligation mixture. Afterward, transformed *E. coli* cells were selected by Ampicilin-supplemented Luria-Bertani medium, then restriction digestion reactions were performed using *Bgl* II and *Kpn* I restriction enzymes to identify the right clones (Fig 3.3 A).

We also constructed promoter containing pGL3-basic and pGL3-enhancer vectors for six different promoter sequences of five different genes (Fig 3.3 B &C).

After screening of the right clone in pGEM-T easy vector using restriction digestion method, digested inserts were extracted from the gel and then inserted into the pGL3-basic and pGL3-enhancer vectors at restriction sites *Bgl II* and *Kpn I* of both vectors. In order to make sure ligation has occurred, *XL-Blue* cells were transformed with the ligation mixture and the ligation mixture was used for transformation in *XL-Blue Ecoli* strain, then the right clones were selected on Ampicillin-supplemented medium and digested with *Bgl* II & *Kpn* I restriction enzymes to check the positive clones and whether they cloned the correct inserts.



1.1kb ladder
2.pGEM-T- *Kcnq1*
3. pGEM-T-*N*ap1l4
4.pGEM-T- *Tssc4*
5.pGEM-T- *Slc22a18*
6.pGEM-T- *Ascl2*

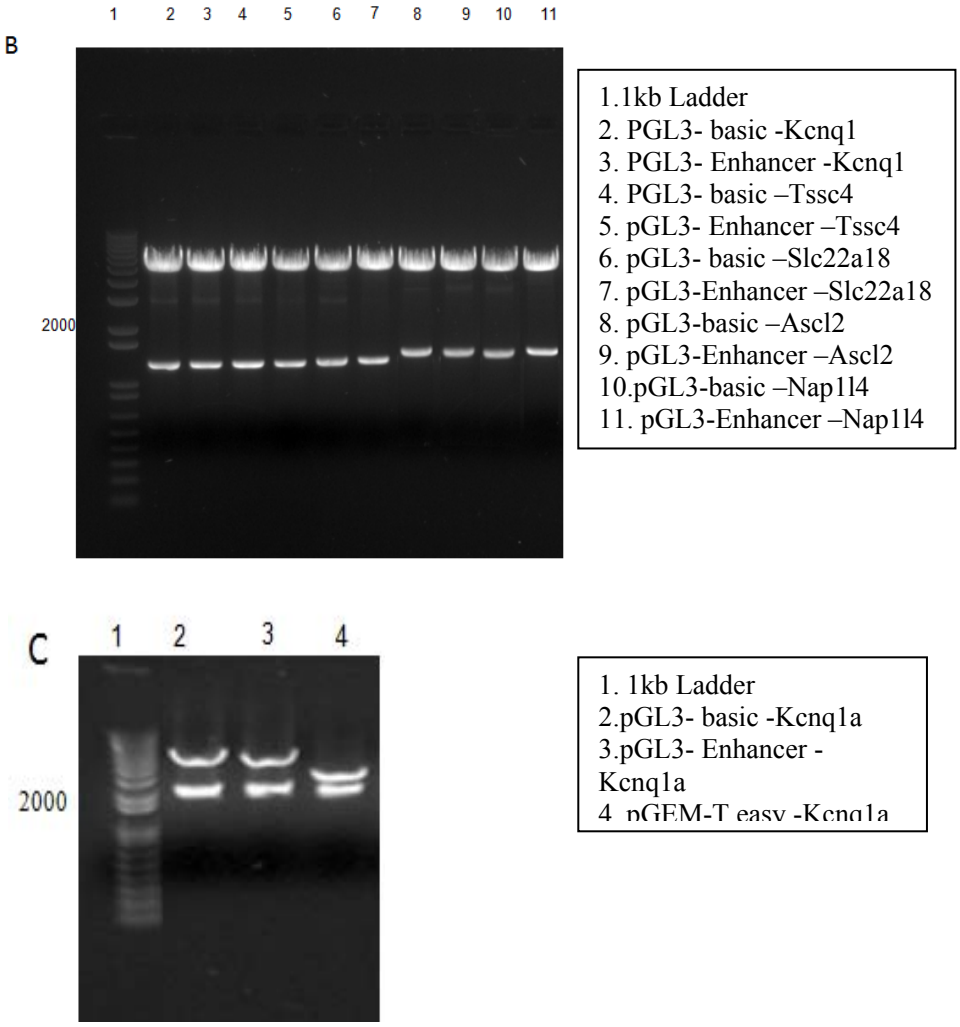


Figure 3.3. Representative results of restriction digestion of three different vectors including pGEM –T easy, pGL3- basic and pGL3-Enhancer with different promoters from *Kcnq1, Kcnq1a, Ascl2, Tssc4, Slc22a18, Nap1l4* and genes which have been loaded next to a 1 kb ladder: A) representative result for promoter regions inserted into pGEM-T easy vector. B) Restriction digestion reaction result for promoters including *Kcnq1*, *Ascl2*, *Tssc4* and *Nap1l4,* which have been cloned into the pGL3, pGL3-basic and pGL3-Enhancer vectors. C) Result for screening of *Kcnq1a* clones using pGEM-T easy, pGL3-basic and pGL3 Enhancer vectors.

3.2 Identification of the different degree of promoter activities

To perform the Luciferase reporter and β-galactosidase enzyme assays, the promoters + luciferase and β-galactosidase reporter constructs were used to assay promoter activities of six promoter sequences from five different genes in *Kcnq1* imprinted locus. Six different promoter fragments were cloned in pGL3-basic and pGL3 –enhancer vector, to drive luciferase and β-galactosidase expression, then the human placental Choriocarcinoma cells JEG3 were transfected with each of these plasmids along with a constant amount of a carrier plasmid PCDNA3 expression vector and pCMV- β-gal vector. The pCMV- β-gal vector was used for normalization of transfection efficiency .

After 48 h, the cell extracts of two identical transfections were analyzed for both luciferase and β-galactosidase enzyme activities using a luminometer and spectrophotometer device, respectively. The measurements were performed in duplicate for each experiment. In addition, the experiment was repeated at least 4 times with two different sets of DNA mini and midi preparations.

All the data from luciferase assay were normalized to the values obtained from β-galactosidase enzyme assay system and plotted in the bar charts. The promoter activities were compared to one obtained using the basic vector and one using pGL3-enhancer vector.

Results of enzymatic assays were evidences for various degrees of promoter activity of the imprinted and non-imprinted genes in the *Kcnq1* locus.

Our data from enzymatic assays indicate that in the absence of enhancer (i.e., when the backbone vector was pGL3-Basic for all six constructs), promoter activities were low compared to one of using pGL3-enhancer vector in JEG-3 cells (Fig 3.4). In addition, promoter strengths of the *Kcnq1* (1.2 kb), *Kcnq1a* (1.9 kb), *Slc22a18, Ascl2, Tssc4* and *Nap1l4* were compared to *Kcnq1ot1* promoter strength as positive control (PC) and a PGL3-basic empty vector (i.e., mock vector) as negative control (NC). The strong promoter activities were observed with Tssc4 placental-specific imprinted gene and Nap1l4 non-imprinted gene, whereas *Kcnq1, Slc22a18* and *Ascl2* showed lower promoter activities.

Besides, the highest activity was seen with *Tssc4* and 1.2 kb *Kcnq1* (ubiquitously imprinted gene) showed lowest activity (Fig 3.4). In addition, the *Kcnq1* Promoter activity was found to be low in both short 1.2kb and long 1.9 kb promoter constructs. However, the later construct showed a major promoter activity in presence of enhancer in PGL3-enhancer vector compared to its counterpart (Fig 3.5). Additionally *Slc22a18* promoter showed almost the same activity as the *Kcnq1* promoter (Fig 3.4).

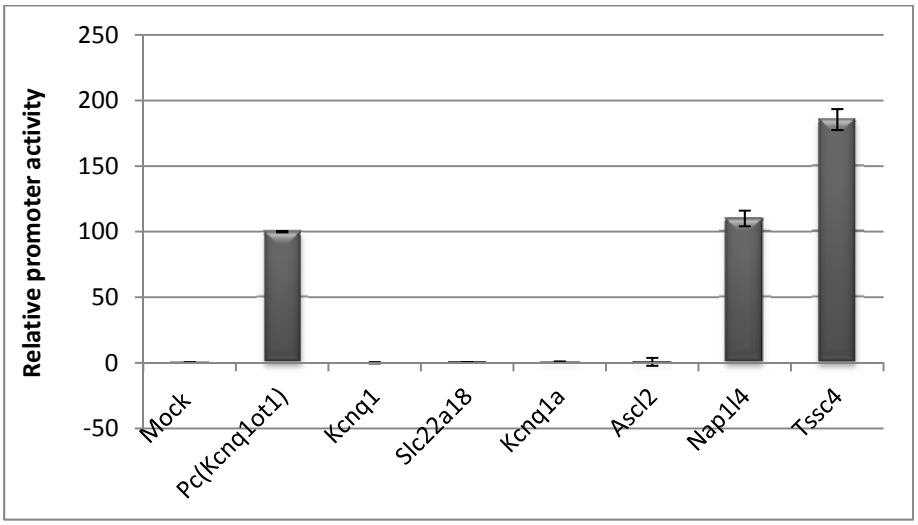


Figure 3.4. Normalized data for promoter activity of the imprinted and non-imprinted genes in the mouse *Kcnq1* locus in the absence of enhancer. Strength of the *Kcnq1* 1.2 kb, *Kcnq1a* 1.9 kb, *Slc22a18*, *Ascl2, Tssc4* and *Nap1l4* promoters were compared to *Kcnq1ot1* promoter strength as a positive control (PC) and pGL3 –basic empty vector (Mock vector) as negative control. The backbone vector was pGL3-Basic for all 6 constructs. The highest activity was observed with *Tssc4* and lowest activity was showed by 1.2 kb *Kcnq1.*

In the other hand, our data from enzymatic assays indicate that in the presence of enhancer (i.e., when the backbone vector was pGL3-enhancer for all six constructs), the promoter activity was found to be higher compared to that of in pGL3-basic vector (Fig 3.5).

Additionally, the highest activity was found with *Nap1l4* (non-imprinted gene) and 1.2 kb *Kcnq1* showed lowest activity. For the rest of constructs different degrees of promoter activities were observed. The promoter strength of *Slc22a18* was very low. Intriguingly, 1.9 kb *Kcnq1a* construct (i.e., longer promoter sequence) showed higher promoter activity

compared to 1.2 kb *kcnq1* (shorter promoter sequence). The placental-specific gene *Ascl2* was found to have a higher promoter activity compared to weaker promoters (i.e., *Kcnq1 and Slc22a18*) (Fig 3.5).

In conclusion, the promoter activity of all six different constructs including *Kcnq1, Kcnq1a, Ascl2, Slc22a18,Tssc4* and *Nap1l4* using pGL3-enhancer vector showed a major increase compared to one of the Basic vector (Fig 3.5). Additionally, the promoter activity of placenta–specific imprinting genes including *Ascl2, Tssc4* and *Nap1l4* non-imprinted genes were observed to be higher than ubiquitously imprinted genes (Fig 3.5).

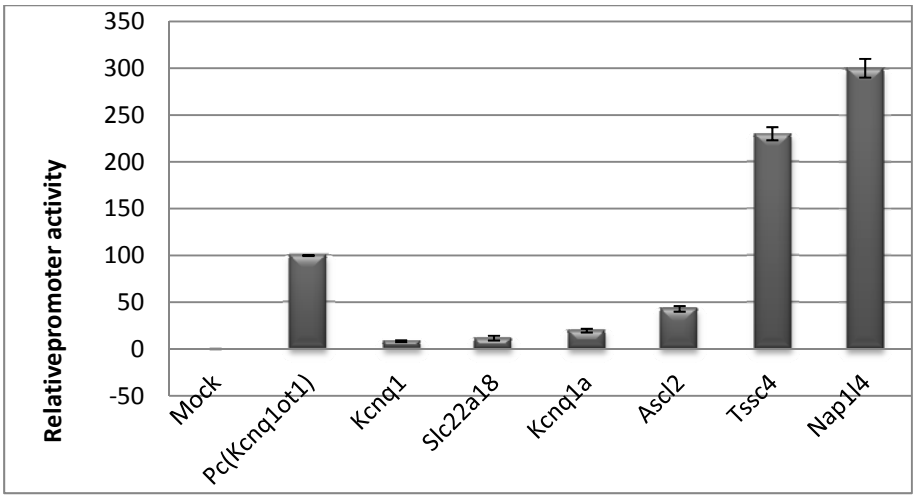


Figure 3.5. Normalized data for promoter activity of the imprinted and non-imprinted genes in the mouse *Kcnq1* locus in the presence of enhancer (i.e., the backbone vector was pGL3-Enhancer for all six different constructs). The promoter strength of the *Kcnq1* 1.2 kb, *Kcnq1a* 1.9kb, *Slc22a18, Ascl2, Tssc4* and *Nap1l4* were compared to *Kcnq1ot11* promoter strength as positive control (PC). The six different constructs including *Kcnq1, Kcnq1a, Tssc4, Ascl2, Nap1l4 and Slc22a18* showed a major improvement in their activities in presence of enhancer. The highest activity was belong to *Nap1l4* and lowest activity was observed by 1.2 *kb Kcnq1*.

4 DISCUSSION

Several studies have been focused on mechanism of imprinting and in particular silencing of the imprinted genes by *Kcnq1ot1*.These studies revealed the fact that this long non-coding RNA can recruit different chromatin modification factors including polycomb repressor complexes and DNA methylases. In addition, it can interact with chromatin region of the promoters in a lineage-specific manner [41].

In this project, we expected to discriminate promoter strength differences that could explain and may uncover one of the multilayer molecular mechanism involved in the regulation of imprinting in *Kcnq1* locus and to find any link between the promoter activity and imprinting. We were also interested to compare promoter activity of the imprinted and non-imprinted genes in *Kcnq1* locus. We were specifically focused on studying the promoter activity of the two ubiquitously imprinted *Kcnq1, Slc22a18* and two placental-specific *Ascl2, Tssc4* genes as well as a non-imprinted *Nap1l4* gene which is located downstream of these imprinted genes in the locus.

It has been shown that, apart from chromosomal organization and nuclear context, other regulatory elements including promoters, enhancers, silencers, locus control regions (LCRs) are key players in transcription initiation and reinitiating. For example, Promoters are involved in initiating transcription [44].

In this project, we found that each individual promoter varies in strength and some of the promoters were enhancer-independent, the promoter activities of placenta specific genes (*Ascl2* and *Tssc4*) either without enhancer or in the absence of this regulatory element were higher than ubiquitously imprinted genes (*Slc22a18* and *Kcnq1*). The *Kcnq1* and *Slc22a18* genes seemed to have weaker promoters. These weak promoters usually have lower rates of transcription than the strong promoters such as *Tssc4* and *Nap1l4*. The *Kcnq1* and *Slc22a18* promoters seem to be more active in the presence of enhancer and an enhancer as a regulatory element can recover these weak promoters. Additionally, the *Nap1l4* and *Tssc4* promoters had higher promoter activities in both conditions (in presence or absence of enhancer) and *Ascl2* observed to have a high degree of promoter activity among the rest of the genes i.e., *Kcnq1, Slc22a18*. Consequently, obtained results may explain why placenta specific imprinted genes are expressed only in embryonic tissue and non-imprinted genes are expressed in both embryonic and extra embryonic tissues.

Our data indicate that degree of promoter strength is involved in imprinting of these genes in the locus and those genes with stronger promoter may escape from imprinting.

In addition, there could be a possible dominant role of strong promoters in regulating neighboring genes with the weak promoters.

It has been reported that promoter-promoter interactions can affect the transcription of mitochondrial genes in *Saccharomyces Cerevisiae* [45]. It has been shown that in a locus, promoters that are in close tandem array can interact and consequently influence the gene expression of one another. In this study, the strong promoters reduce the activity of the weak promoters either in trans or in cis by competing with the weak promoters for RNA polymerase and inhibitory effect of some nucleotides that separating the strong and weak promoter in tandem, respectively [45].

This evidence lets one speculate that in 1 Mb *Kcnq1* locus [41] there is a possibility for a promoter–promoter interaction and strong promoter regulates the weak promoter in the locus. However, further investigations are needed to be done to accept the same type of regulation in *Kcnq1* locus.

Previous studies demonstrated that in the *Kcnq1* locus, those genes located within 200 kb of the *Kcnq1* ICR are ubiquitously imprinted, whereas the placenta specific genes are positioned in a lengthier distance (450 kb) from the *Kcnq1* ICR in a CpG methylation-independent manner [43]. In the other hand, It has been shown that the *Kcnq1ot1* promoter maps to the *Kcnq1* ICR [41].

These data from previous studies were a motivation to propose that, in the *kcnq1* locus those genes located within 200 kb far from the *Kcnq1* ICR/ *Kcnq1ot1* promoter such as *Kcnq1* and *Slc22a18,* which seem to be weaker promoters, are ubiquitously imprinted, whereas stronger promoters of *Tssc4* and *Ascl2* located in 450 kb far from the *Kcnq1* ICR / *Kcnq1ot1* promoter, are able to express only in embryonic tissues.

In the other hand, though *Kcnq1* imprinting domain is a big domain and most of the genes are large, there are possible chances for weak and strong promoters to co close in the domain. To accomplish this cis acting effect in the locus, perhaps function of some unknown elements can facilitate this process.

The data brings about this hypothesis that strong promoter activity of some of the genes and surrounding transcription machinery would be the reason why some of these genes escaping from imprinting. However, further investigations are needed to define a role of promoter strength in imprinting in a parental and tissue specific manners.

5 MATERIALS AND METHODS

5.1 Polymerase chain reaction (PCR)

PCR for 6 promoters was performed in a thermal cycler machine with following reagents: 1 Unit of Taq® DNA polymerase DyNAzyme II 2 U/µl, FINNZYMES, 100-150 ng of genomic DNA, 1x supplied DyNAzyme buffer (FINNZYMES), 1 µl (10 pmol) of both forward and reverse primers, 1 µl of 10 mM dNTPs in a final volume reaction of 25 µl.
The DNA template was initially denaturized at 94° C for 5min, then DNA was amplified by 35-40 cycles of (95° C for 30 s, X° C for 40 s, 72° C for 90 s) and a final extension of 10 minutes at 72° C. The optimized PCR conditions are summarized in (Table 1).

5.2 Gel electrophoresis

In order to make sure that amplification was obtained, gel electrophoresis was performed in a gel unit by using 0.8% and 1% agarose gels in 1X TBE. In order to visualize the PCR products (DNA), 0.2 µg /ml Ethidium bromide was added on to the gels. Then PCR products were loaded into the gel and the electrophoresis was run at 70 ~ / 120 ~ V. Then, PCR products were extracted from the electrophoresis gel using Gel Purification Kit (promega kit) and finally the DNA concentration was measured by nanodrop machine.

Table 1. Primers used for PCR.

Gene	Primer sequence (PCR Conditions 5' →3')		Band size (bp)	Optimized PCR conditions	
				Temp. (°C)	Cycle No.
Kcnq1	Forward	5'ATATggtaccCAGCTGCACAAACCTTCTGA	1203	55°C	35
	Reverse	5'ATATagatctCGGTGAAGTGGTACACGAAA			
Kcnq1a	Forward	5'ATATggtaccATCCTGTTTTCAGCCCACTC	1900	55°C	40
	Reverse	5'ATATagatctAGCTCCAGTGAGAAGGGACA			
Tssc4	Forward	5'ATATggtaccGGCAACCACCACTGTCTCTT	1199	55°C	35
	Reverse	5'ATATagatctCTCAGTAAAGCGCAGCCTCT			
ASCL2	Forward	5'ATATggtaccGGTGCAGTTCATGGCCTACT	1499	57°C	40
	Reverse	5'ATATagatctCTAGCTGGCCTGGAAGTTTG			
Slc22a	Forward	5'ATATggtaccGGGGAAATGGGAAGAGAAAG	1201	55°C	35
	Reverse	5'ATATagatctTTGGGGTAGATTCCACCTTG			
Nap1l4	Forward	5'ATATggtaccCCACATCCACTCCCCAAATA	1444	55°C	40
	Reverse	5'ATATagatctCAACAAAACCACCTGCCATT			

The *Kpn* I (forward primers) and *Bgl* II (reverse primers) restriction sites are in lowercase. The forward and reverse primers respectively contain *Kpn* I and *Bgl* II sites at 5' ends, which were used for cloning in pGl3-basic and pGl3-enhancer vectors.

5.3 pGEM-T easy vector cloning

The cloning of PCR products (i.e., cloning of promoter regions of each gene) into pGEM-T easy Vector was performed in a 10 µl reaction by mixing 5 µl of 2x Rapid ligation buffer (Promega), 1µl of pGEM-T easy vector (50 ng/µl, Promega) and 1 µl of ligase (10 U/µl, Promega) and 6-25 ng of insert. The DNA concentrations of 6 to 25 ng were used for each PCR-products, based on the concentrations obtained from gel-band purification. After preparation of ligation reaction mixtures, they were incubated at room temperature for 2 hours.

5.4 *Escherichia coli* transformation

The transformation reaction was prepared by adding 5 µl of ligation reaction to 50 µl of *E.coli* competent cells (*XL- blue E-coli* strain). The mixture was kept in ice for 30 min and then it was subjected to 42° C heat shock for 90 sec and immediately placed in ice for 3-5 min. After that, 600 µl of Luria-Bertani medium (composition: 10 gm Trypton, 5 gm Yeast extract and 10 gm Sodium chloride in 1 liter sterile H_2O, pH 7.0) was added and the culture was incubated in shaker at 37° C for 1 hour. Later, the transformed reactions were centrifuged at 3000 g for 1 min and plated on to Luria-Bertani (LB) medium containing Ampicillin (100 µg/ml) and incubated at 37° C over night.

5.5 Extraction of plasmid using mini/and midi preparation kit

In order to perform a plasmid extraction, selected transformed bacterial colony was inoculated in 4 ml of LB containing Ampicillin (100 µg/ml) and incubated overnight at 37° C in a shaker with speed of 180-200 rpm and Plasmid was extracted by using Plasmid Qiagen midi and mini DNA prep kit.

5.6 Restriction digestion

Three different digestion reactions were performed in order to digest pGL3-600 basic vector (600 bp insert allow detection of vector digestion with both *Kpn* I and *Bgl* II enzymes), pGL3 –enhancer vector and pGEM-T Easy vector. 20µl reactions were prepared by adding the following reagents: 300-1000ng of the DNA (pGL3-600 basic vector and pGEM-T Easy Vector), 1X digestion buffer (10x concentrated buffer 2, New England Biolabs), 0.5 µl of both *Kpn* I and *Bgl* II enzymes (10 U/µl, New England Biolabs) and 20-(X-3) µl sterile H_2O. These 20 µl reaction mixtures were incubated for 2 to 3 hours at 37° C.

5.7 pGL3-vector (pGL3 basic and enhancer) cloning

The digested inserts from pGEM-T Easy Vectors were cloned into both pGL3 basic and enhancer vectors digested with *Bgl* II and *Kpn* I. The 10 µl ligation reactions were prepared by adding the 1X T4 ligase buffer NEB (10X buffer New England Biolabs), 50ng of pGL3 basic and enhancer Vectors (50 ng/µl promega) and 1 µl of T4 ligase (10 U/µl, NEB) and 30 to 60 ng of the inserts. The amounts of inserts were calculated based on the following formula: Concentration of vector in ng X insert size/vector size X 3 = amount of insert. Finally, all the reactions were incubated at room temperature for 2 hours.

5.8 Cell culture, DNA transfection using calcium chloride ($CaCl_2$) transfection method

JEG-3 a human placental choriocarcinoma cell line was grown in medium containing Minimum essential medium (Eagle), 10% fetal bovine serum (FBS), 2 mM L-Glutamine (5 ml of 200Mm /100X stock solution per 500 ml medium) and 5ml of 100X Penicillin-Streptomycin stock solution per 500 ml (100x stock solution = 10000 U/ml Penicillin G and 10 mg/ml Streptomycin) at 37° C in a sterile chamber with 5% CO2. After preparation of media, in order to prepare a cell density of $1.5X10^6$ per well, the medium, Trypsin and PBS were pre warmed up and cells were washed with PBS, then 1ml of trypsin was added to the cells in order to detaching the cells from T 75 -flask. Later on, 10 ml of the medium was added to the cell suspension and the cells were counted using a hemocytometer. The cell concentration was calculated according to the following formula: Cell concentration per ml = Average number of cells in one large squire (i.e., $1mm^2$) X dilution factor X 10^4. After that, cells were divided into the 6 wells plates about 24 hours before transfection. The 250 µl transfection reactions were prepared by adding the following reagents: 15.5 µl of $CaCl_2$ 2.5M, 108 µl of 10 mM tris pH 7.6 and 250 ng DNA including pGL3-basic-*Kcnq1,* pGL3-basic-*Kcnq1a,* pGL3-basic-*Ascl2,* pGL3-basic-*Slc22a18,* pGL3-basic-*Tssc4,* pGL3-basic-*Nap1l4,* pGL3-Enhancer-*Kcnq1,* pGL3-Enhancer-*Kcnq1a,* pGL3-Enhancer-*Ascl2,* pGL3-Enhancer-*Slc22a18,* pGL3-Enhancer-*Tssc4,*

pGL3-Enhancer-*Nap1l4* (the 250 µl reaction was adapted to the size of the plates) were used for transfection. Additionally, 2.5 µg of a carrier DNA, PCDNA3 vector and 250 ng of pCMV-β-galactosidase were added to all the reactions. 250 ng of pGL3 -742 (contains *Kcnq1ot1* promoter region) and pGL3-Basic vectors (mock vector) as standard positive and negative controls were included. The tubes containing transfected solution were vortexed and 125 µl of 2 X HBS were added drop-wise. The tubes were left at room temperature for 10-20 minutes then the solutions were added drop-wise into the cell plates. Then, plates were incubated at 37° C in 5% CO2, overnight. After 16 hours the cells were washed 3 times with phosphate-buffered saline (PBS) and again fresh medium was added into the plates and incubated.

5.9 Luciferase reporter system and β-galactosidase enzyme assays

At 48 h after transfection, Cells were lysed with 250 µl of 1X Reporter Lysis Buffer (RLB) (Promega). To accomplish the process of lysis, the cells were scraped from the plates and were lysed by a single freeze and thaw. The luciferase assays were performed using a mixture of 100 µl of luciferase reporter assay reagent (Promega) and 20 µl of cell extract. Then the mixture was vortexed and the luciferase readings were performed using a luminometer. Afterwards, Luciferase values were normalized against β-galactosidase enzymatic assay values.

To perform the β-galactosidase enzyme assay, mixtures of 100 µl of cell lysate, 50 µl 1x RLB and 150 µl of 2x β-galactosidase containing buffer were prepared, then the mixtures were vortexed and incubated at 37° C for 30 minutes to develop faint yellow color. Afterwards, the reactions were stopped by adding 500 µl of 1M NaCO3 and instantly the absorbance was read at 420 nm using a spectrophotometer. In each experiment, all the constructs were examined in duplicate. In addition, two different sets of plasmid DNA mini and midi preparations were used for transfection (Qiagen midi or mini DNA prep kit).

6 REFERENCES

1. Pagel, M.D., Pomiankowski, A. (2008) Evolutionary genomics and proteomics .(Evolution of genomic expression. Lemos, B., et al. chapter 5) Sunderland ,Mass: Sinauer Associates Inc. ,U.S.
2. Bonasio, R., et al. (2010) Molecular signals of epigenetic states. Science 330, 612-616
3. Bartolomei, M.S, Ferguson-Smith, A.C. (2011) Mammalian genomic Imprinting. Cold Spring Harb Perspect Biol. 3, a 002592.
4. Waddington, C.H. (1939) An introduction to modern genetics. Allen and Unwin: London.
5. Hurd, P.J. (2010) The era of epigenetics. Brief Funct Genomics 9, 425-428.
6. Hochkiss, R.D. (1948). The quantitative separation of purine, pyrimidines and nucleoside by paper chromatography, J. Biol. Chem. 168, 315–332.
7. Pradhan, S., Esteve, P.O. (2003) Mammalian DNA (cytosine-5) methyltransferases and their expression. Clin Immunol. 109, 6-16.
8. Aapola, U., Liiv, I., and Peterson, P. (2002) Imprinting regulator DNMT3L is a transcriptional repressor associated with histone deacetylase activity. Nucleic Acids Res. 30, 3602-3608.
9. Bock, C., Paulsen, M., Tierling, S., Mikeska, T., Lengauer, T., Walter, J. (2006) CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure. PLoS Genet. 2, e26.
10. Lee, K.K., Workman, J.L. (2007) Histone acetyltransferase complexes: one size doesn't fit all. Nat Rev Mol Cell Biol. 8, 284-295.
11. Dodge J.E., Kang, Y.K., Beppu, H., Lei, H., and Li, E. (2004) Histone H3-K9 methyltransferase ESET is essential for early development. Mol Cell Biol. 24, 2478-2486.
12. Peters, A.H., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., et al. (2003) Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell 12, 1577-1589.
13. Taft, R.J., Pheasant, M., Mattick, J.S. (2007) The relationship between non-protein-coding DNA and eukaryotic complexity. Bioessays 29, 288-99.
14. Prasanth, K.V, Spector, D.L. (2007) Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. Genes Dev. 21, 11-42
15. Ponting, CH.P, Oliver, P.L., Reik, W. (2009) Evolution and functions of long Non-coding RNAs. Cell 136, 629-641
16. Costa, F.F. (2005) Non-coding RNAs: New players in eukaryotic biology. Gene 357, 83-94.
17. Liu, AV.L. Li, y., Tollefsbol, T.O. (2008) Gene-environment interactions and epigenetic basis of human diseases. Mol Biol. 10, 25–36.
18. Swaney, W.T. (2011) Genomic imprinting and mammalian reproduction. Horm Behav. 9, 369–374.
19. Barlow, D.P., Stoger, R., Herrmann, B.G., Saito, K., and Schweifer, N. (1991) The mouseinsulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. Nature 349, 84-872.
20. Peters, J.O., Robson, J.E. (2008) Imprinted non-coding RNAs. Mamm Genome 19, 493–502.
21. Ke, X., et al., (2002) The distinguishing sequence characteristics of mouse imprinted genes. Mamm Genome 13, 639-45.
22. Temple,,K., Shield, J.P.H. (2010) 6q24 transient neonatal diabetes. Rev Endocr Metab Disord. 11, 199–204.
23. Bastepe , M., Altug-Teber, O., Agarwal, C., Oberfield, SE., Bonin, M., Jüppner, H. (2011) Paternal uniparental isodisomy of the entire chromosome 20 as a molecular cause of pseudohypoparathyroidism type Ib (PHP-Ib). Bone 48, 659–662.
24. Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M.A. (2002) Epigenetic reprogramming in mouse primordial germ cells. Mech Dev. 117, 15-23.
25. Allegrucci, C., Thurston, A., Lucas, A., Young, L. (2005) Epigenetics and the germline. Reproduction 129, 137–149.

26. Arnoud, PH. (2010) Genomic imprinting in germ cells: imprints are under control. Reproduction 140, 411–423.
27. Ooi ,S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin S.P., Allis C.D., et al. (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 448, 714–717.
28. Chotalia, M., Smallwood, S.A., Ruf, N., Dawson, C., Lucifero, D., Frontera, M., James, K., Dean, W., & Kelsey, G. (2009) Transcription is required for establishment of germline methylation marks at imprinted genes. Genes Deve. 23,105–117.
29. Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R. & Sasaki, H. (2008) Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. Genes Deve. 22, 1607–1616
30. Bartolomei, M.S. (2009) Genomic imprinting: employing and avoiding epigenetic processes. Genes Dev. 23, 2124-2133.
31. Handoko, L., et al. (2011) CTCF-mediated functional chromatin interactome in pluripotent cells. Nat genetics 10,1038/ng.857.
32. Arney, K.L., Bae, B., Olsen, C., Drewell, R.A. (2006) The human and mouse H19 imprinting control regions harbor an evolutionarily conserved silencer element that functions on transgenes in Drosophila. Dev Genes Evol. 216, 811–819.
33. Vu, T.H., Li, T., Hoffman, A.R. (2004) Promoter-restricted histone code, not the differentially methylated DNA regions or antisense transcripts, marks the imprinting status of IGF2R in human and mouse. Hum Mol Genet. 13, 2233-2245.
34. Mercer, T.M., Dinger, M.E., and Mattick, J.S. (2009) Long non-coding RNAs: insights into functions. Nat Rev Genetics 10, 155-159.
35. Ling Chen, L., and Carmichael, G.G. (2010) Decoding the function of nuclear long non-coding RNAs. Cell Biol. 22, 357–364.
36. Rinn, J.L., Kertesz , M., Wang , J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., GoodnOugh, L.H., Helms, J.A., Farnham, P.J., Segal, E., Chng, H.Y. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by non-coding RNAs. Cell 129, 1311–1323.
37. Bond, C.S., Fox, A.H. (2009) Paraspeckles: nuclear bodies built on long non-coding RNA. Cell Biol. 186 , 637-44.
38. Prasanth, K.V., Prasanth, S.G., Xuan, Z., Hearn, S., Freier, S.M., Bennett, C.F., Zhang, M.Q., Spector, D.L. (2005) Regulating gene expression through RNA nuclear retention. Cell 123, 249-63.
39. Mohammad, F., Mondal, T., Kanduri, CH. (2009) Epigenetic of imprinted long non-coding RNAs. Epigenetics 4, 277-86
40. Mohammad, F., Mondal, T., Guseva, N., Kumar Pandey, G., Kanduri, CH.(2010) Kcnq1ot1 non-coding RNA mediates transcriptional gene silencing by interacting with Dnmt1. Development 137, 2493-9
41. Kanduri, C.H., (2011) *Kcnq1ot1*: A chromatin regulatory RNA. Seminars in Cell Deve Biol. 22, 343– 350
42. Pandey, R.R., Ceribelli, M., Singh, P.B., Ericsson, J., Mantovani, R., Kanduri, CH. (2004) NF-Y regulates the antisense promoter, bidirectional silencing, and differential epigenetic marks of the *Kcnq1* imprinting control region. J Biol Chem. 279, 52685–93.
43. Pandey, R.R., Mondal. T., Mohammad. F., and Kanduri, C. (2008) Kcnq1ot1 antisense non-coding RNA mediates lineage specific silencing through chromatin-level regulation. Mol Cell 32, 232-46.
44. Burns, J.L., Jackson, D.A., and Bassim Hassan, A. (2001) A view through the clouds of imprinting. FASEB. 15, 1694-1703
45. Biswas, T.K. and Getz, G.S. (1988) Promoter-promoter interactions influencing transcription of the yeast mitochondrial gene, Oli 1, coding for ATPase subunit 9 : cis and trans. J. Biol. Chem. 263, 4844–4851

7 ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor Prof. Chandrasekhar Kanduri for his inspiring guidance, continues encouragement and given me freedom in carrying out this project.

My special thanks to Dr. Volkmar Passoth, for his support on my registration and official regulation, and also his knowledge, patience and advises to complete this thesis.

I owe my sincere thanks to Gaurav Pandey Kumar and Tanmoy Mondal for their constant support and their advises to carry out the lab work.

I would be failed in my duties, if I don't appreciate the help, support and encouragement, which I received from my family, my special thanks to my husband Farzad Bakhtiyari and my son Sourena.

Finally, I must thank to all of those who gave me the possibility to complete this thesis.